

The College Readiness Data Catalog Tool: User Guide



Tools

Sheila M. Rodriguez
Education Development Center, Inc.

Angela Estacion
Quill Research Associates, LLC

A data catalog can help education leaders and researchers determine whether sufficient data exist to undertake a research study. In 2013 the Regional Educational Laboratory Northeast & Islands created a data catalog to help the U.S. Virgin Islands College and Career Readiness Alliance assess the feasibility of a study to identify the strongest indicators of college readiness in the U.S. Virgin Islands. The resulting College Readiness Data Catalog Tool is a flexible-use Excel workbook that provides a shell for organizing and tracking student data relevant for measuring college readiness. This tool can help states and districts create similar data catalogs to inventory their data systems, assess the availability of college readiness indicators, and identify gaps. Researchers can use the resulting data catalogs to determine the feasibility of college readiness-related studies.

This user guide describes the College Readiness Data Catalog Tool, which was developed by the Regional Educational Laboratory Northeast & Islands (REL-NEI) in response to a request from one of its research partners, the U.S. Virgin Islands College and Career Readiness Alliance (USVI Alliance). This alliance of state, district, and community education leaders aims to help the U.S. Virgin Islands Department of Education build a longitudinal data system for preschool through college and incorporate a college and career readiness indicator system.

The USVI Alliance wanted to identify the strongest indicators of college readiness in the U.S. Virgin Islands. To determine whether sufficient data exist for a research study on this topic, REL-NEI researchers developed the College Readiness Data Catalog Tool, which can be used to compile a research-based data catalog and summary report to assess a study's feasibility or refine its research questions.¹ The amount of time needed to complete a data catalog depends on the amount of information available and on the number of people and agencies involved.²

Purpose of this tool

As the name indicates, the College Readiness Data Catalog Tool focuses on identifying data that can indicate a student's college readiness. While college readiness indicators may also signal career readiness, many states, districts, and other entities, including the U.S. Virgin Islands, do not systematically collect career readiness indicators.

Although the College Readiness Data Catalog Tool was developed with the needs of the USVI Alliance in mind, it is flexible enough to be used by other interested parties. This tool allows states, districts, and other entities to create data catalogs to assess the availability of college readiness indicators and identify gaps that may present challenges for indicator systems. Researchers may also use data catalog tools as they work with states, districts, and other entities to determine the feasibility of studies on college readiness, including, for example, identifying the strongest indicators in a given context. The tool is meant to be used only after a research question has been identified.

The College Readiness Data Catalog Tool is a flexible-use Excel workbook that provides a shell for organizing and tracking student data relevant for measuring college readiness. The shell organizes data at three levels:

- *Constructs*. Concepts identified during the literature scan that are of interest for measuring. Broad and indirectly measurable, constructs identify important ideas that should be measured to answer each research question and meaningfully group the information in the catalog. College readiness constructs are overarching categories that are highly relevant for predicting success in college.
- *Indicators*. Concepts that provide greater specificity than constructs. Not all constructs have associated indicators, and some have more than one.
- *Data elements*. Specific data sources that describe specific and measurable concepts. Data elements include traditional sources such as measures from assessments, performance evaluations, and formal observations, as well as documents such as emails, agendas, and meeting notes.

The tool also provides a list of key data element characteristics, such as years of available data, information about linking, changes to data elements across years, and specific fields or values. The data element characteristics collected may vary. For example, this tool was meant to inventory only student-level information. Therefore, data level (student, classroom, teacher, school district) was not relevant and not included. However, the sample data catalog summary report (appendix A) presents a data catalog that required collecting this data element characteristic. All three data levels (construct, indicator, and data element) and their characteristics are important for informing decisions about using college readiness data.

Materials included in this user guide

This user guide highlights the purpose of data catalogs and describes how states, districts, or schools might best use them. It explains how to use the tool to inventory college readiness data. Step-by-step instructions are provided, as well as a sample data catalog summary report (appendix A) and a template for a data catalog summary report (appendix B). The tool itself (a Microsoft Excel file) is available for download from the REL-NEI website (<http://www.relnei.org>) or by contacting REL-NEI at relneiinfo@edc.org.

Tool components and organization

REL-NEI researchers used the National High School Center's College and Career Development Organizer to guide the constructs and examples of indicators (National High School Center, 2012). The center's work, summarizing a diverse body of research, policy, and advocacy documents, outlines the skills and knowledge

deemed necessary to be college and career ready and organizes them into three components: core content, pathways content, and lifelong learning skills. The center also developed a framework to measure these skills and knowledge through several constructs, three of which were used in the current tool: high school academic performance and engagement, postsecondary access and enrollment, and attainment and authentication: postsecondary certification and success.

Within these constructs are associated indicators of high school graduation and college success. Each indicator includes characteristics that are necessary to answer potential research questions about college readiness that will guide states and districts in developing a robust catalog of readiness indicators. The list is not meant to be exhaustive, but to provide a range of examples. Users may include more or fewer elements.

Construct 1. High school academic performance and engagement

This construct includes outcomes and measures identified in the literature as indicators of students on track for high school graduation. These indicators may also identify high school students as “college ready.”

This construct includes student- and school-level indicators on:

- Enrollment status (beginning-of-year status, end-of-year status, dates of enrollment).
- Demographic data (number of students, number of English language learner students, students by race/ethnicity).
- Course performance and data on credit accumulation and recovery (grades in every subject, average grades for each subject).
- Grade point average (calculation of academic achievement, usually given on a 4.0 scale).
- Performance on high school end-of-course or exit exams (final scores or average scores).
- SAT and ACT scores (scale scores for SAT, composite score for ACT).
- Attendance (absences, tardiness).
- Suspensions/behavior factors (number of suspensions, disciplinary action).
- High school credential (high school diploma, GED).
- Performance on aligned assessments of high school core content and college placement tests (for example, Partnership for Assessment of Readiness for College and Careers [PARCC] summative and nonsummative assessments, ACT Compass placement test).
- Participation in accelerated learning programs and college-ready courses (for example, Advanced Placement, International Baccalaureate coursework).
- Postsecondary credit accumulation and recovery, per discipline; college credit earned prior to postsecondary enrollment (for example, dual enrollment or dual credit courses, credit recovery classes).
- Responses on social and emotional learning surveys (for example, the Comprehensive School Climate Inventory).

Construct 2. Postsecondary access and enrollment

This construct includes steps in the application process that may be necessary for eventual college success.

This construct includes indicators on:

- Free Application for Federal Student Aid (FAFSA) completion (completing the FAFSA online application).
- Postsecondary applications completion (completing applications to community colleges, two- and four- year colleges, and universities or technical colleges).
- Postsecondary program enrollment (getting in, paying initial registration fees, and signing agreements to attend).

Construct 3. Attainment and authentication: postsecondary certification and success

This construct includes indicators of attaining a degree or certificate at a postsecondary institution. The National Student Clearinghouse (NSC) of postsecondary enrollment and attainment is the most comprehensive national data source to identify college enrollment and degree completion. Ninety-four percent of students attending public and private U.S. colleges and universities in 2011 were represented in the NSC database (NSC, 2012). Entities may be able to link their own data to NSC data by student name, date of birth, and high school attended.

This construct includes indicators on:

- Postsecondary enrollment (enrollment at community college, technical college, or two- and four-year college or university).
- Postsecondary course performance (earned credits, grade point averages).
- Postsecondary degrees (degrees earned at a postsecondary institution, such as associate's, bachelor's, or technical degree).

Template for a summary report

REL-NEI researchers completed a data catalog using a tool similar to the College Readiness Data Catalog Tool described in this user guide. They also produced a summary report for the USVI Alliance that describes the background and intent of the data catalog, identifies gaps in the data, and discusses recommendations for study feasibility and limitations to consider. Any entity using the tool should create a similar summary report to provide an overview of the project and findings. This summary report may be useful for briefing others on the status of data elements and the implications of any weaknesses in data collection or quality.

This user guide provides a sample summary report (appendix A) and a template for a summary report (appendix B). Because REL-NEI agreed not to make the USVI data catalog publicly available, and no additional college and career readiness-related data catalogs were completed, the sample summary report presents the results from a data catalog of programs and services for English language learner students in one masked district.

Using the tool

To use this tool, you will need a PC computer with Microsoft Excel version 2003 or later or a Mac computer with Excel version 11.0 or later.

Opening the tool

To open the tool, double-click the file icon or name where it is saved. Alternatively, you can open the file within Excel by clicking Open on the File tab. After locating the file, click the Open button.

When you open the tool, you will see the preloaded research question on the top left (figure 1). The constructs are listed directly below. The subsequent column lists the indicators within each construct, followed by another column that will contain the data elements pertaining to each indicator.

As you scroll to the right, you will see columns containing various data element characteristics. You can also use the Tab button to move between columns. To enter information about each data element characteristic, click on the appropriate cell and begin to type.

Figure 1. Partial data catalog tool

Identifying data elements and creating new rows

In the data catalog tool, each construct has a list of relevant indicators. In turn, each indicator provides space for up to three data elements. Review your datasets, codebooks, or data sources—including traditional sources such as assessments, performance evaluations, and formal observations—to identify the data elements that match these indicators and constructs. You may create additional rows to include more than three data elements for a given indicator. Remember to rename your file and save it to your computer.

Entering data element characteristics

After identifying the final list of data elements, you will describe important characteristics related to each element. Data element characteristics to be collected, when available, include:

- **Years.** Identifies the school years for which data are available or the number of years the data are available for each element.
- **Can the data element be linked?** Describes whether the specific data element is linkable to other data elements (for example, whether students' high school data elements are linked to their postsecondary data elements).
- **Changes to data.** Lists any changes to the data sources that measure the construct and any changes to the data over time or across populations.
- **Response categories.** Lists the format of the data element.
- **Missing data.** Identifies any missing data needed to answer the research question.
- **Known misclassifications.** Identifies any entry errors or inconsistencies in the data elements.
- **Audit information.** Provides additional information on improving the data.
- **Why was this element collected? How will it be used?** Provides background information about the data collection process and purposes.
- **Anticipated changes to the data element.** Lists any changes that may occur for the data element.
- **Notes.** Includes any additional information that may be important when planning to use the data element. Not all data elements will have notes.

In a completed data catalog, each data element has its own row for describing the data element characteristics. Some data elements may have the same characteristics because they come from the same source, resulting in repetitive text. In some situations you may want to repeat text for each data element to produce a complete inventory of all available data elements.

Appendix A. Sample data catalog summary report

This appendix provides an example of a data catalog summary report, constructed using the template in appendix B. This example is based on English language learner student population and not college readiness, so it does not match the discussion of the College Readiness Data Catalog Tool. It is simply intended to show what a summary report could entail.

Description of project

Through the English Language Learner Alliance, the Regional Educational Laboratory Northeast & Islands (REL-NEI) is collaborating with states, districts, and schools on using data to understand the diversity of their English language learner student populations and explore the programs and services that best meet these learners' diverse needs. Through REL meetings and activities, alliance and REL staff developed the following research questions:

- What are the characteristics of teachers providing instruction to English language learner students?
- What is the preparation of teachers of English language learner students?
- What are the characteristics of instructional programs for English language learner students?

REL staff developed a data catalog to determine whether sufficient data were available to answer these questions and, if not, to recommend modifications to the research questions to meet alliance needs. This summary and the attached data catalog document the available data, identify data gaps, and serve as a starting point for research studies to answer the alliance research questions.

REL staff worked with alliance members from one of the REL-NEI Region states to identify and collect information on the data available in either electronic or paper format. Information from the following documents is included in the data catalog:

- A 2010/11 and 2011/12 desk audit.³
- Annual permission slips, which provide parental approval for student participation in English language learner programs.
- State department of education desk audit, Chapter 16–54/Title III district monitoring checklist for 2009/10.
- State department of education English language learner student identification procedures.
- English language learner student exit criteria checklist and sign-off form.
- State department of education home language survey.
- Individual Mainstreamed Student Monitoring Report from local school department English language learner programs (a progress report that is attached to students' report cards in the district).
- Documentation of the data elements included in the local district personnel data system.
- District English language learner program description worksheet, which describes the English language learner programs in detail and is sent to the state department of education annually.
- District-developed document that describes English language learner instructional program models available in district schools.

A district liaison sent an additional document that showed district purchases. REL-NEI staff who reviewed the information determined that there was insufficient detail to document supplemental materials purchased to support instruction for English language learner students. Therefore, the catalog does not include any information from the purchasing database.

Data catalog format and key constructs

Below are the data catalog format and key constructs for the data catalog. The x-axis of the catalog includes the elements described below. The letter corresponds to the column label in the Excel version of the catalog.

- A. *Research question*. Identifies the research question to which the information in that row pertains.
- B. *Construct*. Identifies concepts found during the literature scan that are of interest for measuring. Broad and indirectly measurable, the constructs are used to identify important ideas that should be measured to answer each research question and meaningfully group the information within the catalog.
- C. *Indicators*. Provide greater specificity than constructs. Not all constructs have associated indicators.
- D. *Data elements*. Describes specific data sources that measure the construct. Data elements include traditional sources such as measures from assessments, performance evaluations, and formal observations, as well as documents such as emails, agendas, and meeting notes.
- E. *Data source* (for example, database, table, template, text-based document). Identifies where the data element is located.
- F. *Responsible office*. Identifies the district department office responsible for housing the data element.
- G. *Years*. Identifies the school years for which data are available.
- H. *Level of data for linking*. Provides information about how data elements can be connected, such as to students, teachers, or schools.
- I. *Response categories*. Lists the format of the data element, where available.
- J. *Reason for collecting and using the element*. Provides background information about the data collection process and purposes.
- K. *General notes*. Includes any additional information that may be important when planning to use the data element. Not filled in for all data elements.

Several constructs that were important for addressing one or more of the research questions align with the state department of education data collection about English language learner programs. English Language Learner Alliance constructs are reflected in the second column of the data catalog, with several data elements available to measure each construct. The data catalog includes the following constructs:

- *Program model*. Describes instructional approach used with English language learner students.
- *Program goals*. Defines language and academic goals associated with instructional programs.
- *Primary program features*. Describes features of each component of the instructional delivery and support programs, such as academic content, assessments, and accommodations.
- *Parental involvement*. Describes the districts' approach and structure for encouraging parent involvement in children's education.
- *Target population*. Describes student academic and background characteristics, primarily to identify and place students appropriately.
- *Teacher characteristics*. Describes teacher background characteristics and employment status.
- *Teacher preparation*. Includes information about staff training, qualifications, and professional development opportunities.

Gap analysis

The following information describes gaps in the data available to answer each of the three research questions.

What are the characteristics of teachers providing instruction to English language learner students?
The human resources database includes teacher information such as salaries, demographics, certifications, local experience, endorsements and licenses, and whether the teachers are certified to teach English language learner students. Teacher characteristics beyond the database information were not identified.

What is the preparation of teachers of English language learner students? REL-NEI staff and the district liaison agreed that answering this question required capturing teachers' preparation for working with English language learner students before teaching in the district (or before teaching English language learner students in the district) as well as their preparation while employed in the local school district.

The local school district offers a wide variety of training and professional development programs for English language learner teachers, and information about teacher participation is captured in an electronic data file. The district liaison was unable to provide this documentation from the system because of changes in central office staff responsible for documenting teacher participation in the districts' training and professional development opportunities. However, according to the district liaison, the system captures the following teacher information: each course, workshop, or professional development offering attended; dates of attendance; and the number of hours of participation, measured in continuing education unit equivalents. The system also identifies whether the teacher is certified to teach English language learner students. The system does not capture college courses or other professional development programs that are offered outside the district.

The district stores data on licenses, endorsements, and certification, including expiration dates, in a human resources database. The system does not capture detailed data about teachers' pre-hire preparation, such as test scores, academic transcripts, or recertification courses.

What are the characteristics of instructional programs for English language learner students? The local school district maintains a substantial amount of data that can answer this question. The data describe program goals, delivery models, instructional methods, assessments used, and other characteristics. Much of the local school district information is prepared for submission to the state department of education and is therefore available in all of the state's school districts. The information could be used to conduct a descriptive study of English language learner programs throughout the state. Data from the state department of education's annual census for English language learner students links students to instructional program delivery models and characteristics.

One gap in the data on instructional programs concerns the curriculum and supplemental materials. The local school district requires all students to use the same core curriculum materials, regardless of their English language learner status. However, the English language learner office purchases supplemental materials to support students' ability to access the curriculum and master academic content. There is no documentation of supplemental materials purchased.

Other important limitations

One of the main data sources reviewed, the district English Language Learner Program Description Worksheet, was prepared for one-time submission to the state department of education based on data for the 2010/11 school year. In the local school district there were no major changes to English language learner program delivery models since the submission. These data are available to all school districts in the state. However, if researchers use these data, they need to determine whether there were substantive changes to program delivery models in each district before drawing conclusions relevant to current programs.

Recommendations for study feasibility

In collaboration with the state department of education, REL-NEI researchers could pursue a study describing English language learner program delivery models statewide in the 2010/11 school year. Before conducting such a study, researchers would need to determine whether substantive changes were made to the

English language learner program delivery models since the one-time data submission to the state department of education. If a substantial proportion of districts have changed their delivery models, it is unclear how useful the information would be to alliance members.

In the local school district there were insufficient data to describe teachers' preparation before they were employed in that district or their post-employment preparation and professional development activities outside of district-sponsored opportunities. If this is a common finding across alliance districts and states, REL-NEI and the alliance could discuss whether REL-NEI should identify critical data for districts to collect about teacher preparation and external training and professional development experiences, develop a survey or other data collection instrument that would permit districts to collect these data, and support analysis of district-collected data.

The state department of education has begun collecting data to link students to teachers and the courses they teach. While the first year of state data collection is not likely to produce reliable (or complete) information, this collection could eventually become a rich data source for researchers. Combining these data with information on program delivery models and licensing data collected by the state may provide the alliance with robust information to answer research questions related to the impact on student outcomes of program delivery models and teacher preparation. A first step in understanding the potential value of these data would be to conduct a more detailed analysis or audit of the newly collected data with a targeted research perspective and a closer look at key data sources.

Appendix B. Template for a data catalog summary report

This template for a data catalog summary report lays out what should be included in a summary report. This is useful for briefing others on the status of data elements and the implications of any weaknesses in data collection or quality. Once you have identified a research question, you may use this data catalog tool to develop a research-based data catalog and a summary report to help determine whether sufficient data exist to answer the research question.

Description of project

Provide a two- to three-paragraph description of the research project, identifying the need for the project, research questions guiding the project, and the data sources considered.

Data catalog format and key constructs

Describe the data catalog tool, explaining the column headings (that is, the types of information to be collected in the data catalog) and discussing the constructs, indicators, and sample data elements included.

Gap analysis

Summarize the fit between constructs and data elements. Discuss any constructs with missing or poorly fitting elements.

Other important limitations

Summarize any other data limitations that may affect the study. For example, does the complete set of data elements needed exist for only certain years? Was there a change in the data collection process that falls within the years of study?

Recommendations for study feasibility

Make a recommendation for study feasibility and include any suggested changes to the study design or research question wording given the data.

Notes

1. For this tool the team focused on one research question, which aligns with the preloaded constructs and indicators; other data catalogs may include more than one research question per catalog. The constructs and indicators should align with the research questions.
2. For example, completing a data catalog with poorly documented data from multiple agencies will take much longer than completing a data catalog with well documented data from a single source (for example, codebooks, historical personal knowledge).
3. The desk audit was a process in which the research team identified and reviewed hard copies of active files on people involved in relevant professional development activities. For instance, professional development activity sign-in sheets provided the only record of participation in certain activities.

References

National High School Center. (2012). *College and career development organizer*. Washington, DC: Author.
<http://eric.ed.gov/?id=ED532567>

National Student Clearinghouse. (2012). *Enrollment coverage worksheet*. Retrieved from <http://nscresearchcenter.org/workingwithourdata/>

The National Center for Education Evaluation and Regional Assistance (NCEE) conducts unbiased large-scale evaluations of education programs and practices supported by federal funds, provides research-based technical assistance to educators and policymakers, and supports the synthesis and widespread dissemination of the results of research and evaluation throughout the United States.

September 2014

This report was prepared for the Institute of Education Sciences (IES) under Contract ED-IES-12-C-009 by Regional Educational Laboratory Northeast & Islands administered by Education Development Center, Inc. The content of the publication does not necessarily reflect the views or policies of IES or the U.S. Department of Education nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

This REL report is in the public domain. While permission to reprint this publication is not necessary, it should be cited as:

Rodriguez, S. M., and Estacion, A. (2014). *College Readiness Data Catalog Tool: User guide*. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast & Islands. Retrieved from <http://ies.ed.gov/ncee/edlabs>.

This report is available on the Regional Educational Laboratory website at <http://ies.ed.gov/ncee/edlabs>.

The Regional Educational Laboratory Program produces 7 types of reports

	Making Connections Studies of correlational relationships
	Making an Impact Studies of cause and effect
	What's Happening Descriptions of policies, programs, implementation status, or data trends
	What's Known Summaries of previous research
	Stated Briefly Summaries of research findings for specific audiences
	Applied Research Methods Research methods for educational settings
	Tools Help for planning, gathering, analyzing, or reporting data or research